

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:29 ; Search time 299.73 seconds  
(without alignments)  
23.463 Million cell updates/sec

Title: US-09-331-631a-7\_COPY\_81\_140  
Perfect score: 342  
Sequence: 1 LORQYQCCGRCQEQDQGOR.....HENYHNHKKNRSEEGQOR 60

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP:REML\_15:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	100.0	525	10 Q43358	Q43358 theobroma c
2	143.5	42.0	593	10 Q9SEW4	Q9SEW4 juglans reg
3	130.5	38.2	666	10 Q9SP14	Q9SP14 macadamia i
4	127.5	37.3	625	10 Q9SP13	Q9SP13 macadamia i
5	126.5	37.0	666	10 Q9SP15	Q9SP15 macadamia i
6	113	33.0	2123	5 Q909S7	Q909S7 dictyostell
7	104	30.4	810	10 Q9ZW13	Q9ZW13 cucurbita m
8	103.5	30.3	1038	5 Q60983	Q60983 dictyostell
9	101.5	29.7	304	4 Q15409	Q15409 homo sapien
10	100.5	29.4	1556	5 Q906A1	Q906A1 drosophila
11	100	29.2	523	5 Q901K4	Q901K4 drosophila
12	100	29.2	666	5 Q9W318	Q9W318 drosophila
13	96	28.1	171	11 Q61118	Q61118 mus musculu
14	94.5	27.6	420	11 Q35861	Q35861 mus musculu
15	93.5	27.3	828	5 Q9VZU6	Q9VZU6 drosophila
16	93	27.2	930	10 Q9SZY9	Q9SZY9 arabidopsis
17	92.5	27.0	838	5 Q9Y0C9	Q9Y0C9 dictyostell
18	92	26.9	139	11 Q62006	Q62006 mus musculu
19	92	26.9	454	5 P91642	P91642 drosophila

20	92	26.9	531	5 Q9VT57	Q9VT57 drosophila
21	92	26.9	2074	11 Q88542	Q88542 mus musculu
22	91.5	26.8	517	5 Q9W4E8	Q9W4E8 drosophila
23	91	26.6	1457	5 Q44011	Q44011 dictyostell
24	91	26.6	3502	5 Q9VJ9	Q9VJ9 drosophila
25	90.5	26.5	776	5 Q9VFR7	Q9VFR7 drosophila
26	89	26.3	1094	5 Q23915	Q23915 dictyostell
27	89	26.0	259	5 Q9W483	Q9W483 drosophila
28	89	26.0	406	2 Q87306	Q87306 borrelia bu
29	89	26.0	646	2 Q9KW25	Q9KW25 vibrio chol
30	89	26.0	821	5 Q9W3A1	Q9W3A1 drosophila
31	89	26.0	853	4 Q9NTG8	Q9NTG8 homo sapien
32	89	26.0	1010	5 Q18346	Q18346 drosophila
33	89	26.0	2678	5 Q9NDS4	Q9NDS4 dictyostell
34	88.5	25.9	1326	5 Q9W4N4	Q9W4N4 drosophila
35	88	25.7	242	10 Q81251	Q81251 zea mays su
36	88	25.7	242	10 Q9SBE6	Q9SBE6 zea mays su
37	88	25.7	697	5 Q9VZX2	Q9VZX2 drosophila
38	88	25.7	791	5 Q00841	Q00841 dictyostell
39	88	25.7	1502	4 Q9ULJ8	Q9ULJ8 homo sapien
40	88	25.7	2150	5 Q23863	Q23863 dictyostell
41	87	25.4	420	5 Q76853	Q76853 dictyostell
42	87	25.4	551	10 Q43607	Q43607 pinus dulc
43	86.5	25.3	57	4 Q9Y4J1	Q9Y4J1 homo sapien
44	86.5	25.3	341	5 Q9VR72	Q9VR72 drosophila
45	86.5	25.3	438	5 Q9W2K6	Q9W2K6 drosophila

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	525 AA.
ID	Q43358			
AC	Q43358			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DE	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VICILIN PRECURSOR.			
GN	CSV.			
OS	Theobroma cacao (Cacao).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;			
OC	Malvales; Malvaceae; Theobroma.			
OX	NCBI_TaxID=3641;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LEAVES:			
RX	MEDLINE-92288309; PubMed-1600151;			
RA	McHenry L., Fritiz P.J.;			
RT	"Comparison of the structure and nucleotide sequences of vicilin genes			
RL	of cocoa and cotton raise questions about vicilin evolution.";			
RL	Plant Mol. Biol. 18:1173-1176(1992).			
DR	EMBL; X62625; CAA44493.1; -			
DR	EMBL; X62626; CAA44494.1; -			
DR	HSSP; P02853; 2PHL.			
DR	MENDEL; 30919; TheCG:1188;30919.			
DR	INTERPRO: IPR001113; -			
DR	PFAM; PF00546; Seedstore-7s; 1.			
DR	PRODOM; PD081059; -; 1.			
FT	SIGNAL.			
FT	POTENTIAL.			
FT	CHAIN 25 525 VICILIN.			
SQ	SEQUENCE 525 AA; 60798 MW; 19114C5C248905D CRC64;			

Query Match 100.0%; Score 342; DB 10; Length 525;  
Best Local Similarity 100.0%; Pred. No. 2.2e-30;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LORQYQCCGRCQEQDQGORCKCWEQKEDRGHEHYNHKKNRSEEGQOR 60  
DB 81 LORQYQCCGRCQEQDQGORCKCWEQKEDRGHEHYNHKKNRSEEGQOR 140

RESULT	2			
09SEW4				
ID	09SEW4	PRELIMINARY:	PRT;	593 AA.
AC	09SEW4;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, last annotation update)			
DE	VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).			
OS	Juglans regia (English walnut).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosid1 I;			
OC	Fagales; Juglandaceae; Juglans.			
OX	NCBI_TaxID=51240;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. SUNLAND. TISSUE=SOMATIC EMBRYO LINE;			
RA	Teuber S.S., Jarvis K.C., Peterson W.R., Danekar A.M., Ansari A.A.;			
RT	"Identification and cloning of a cDNA encoding a vicilin-like protein			
RT	Jug r 2, from English walnut kernel (Juglans regia): a major food			
RT	allergen ";			
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF066055; AAF18269.1; .			
DR	HSSP; P02853; 2PHT.			
DR	INTERPRO; IPR001113; .			
DR	PFAM; PF00546; Seedstore_7s; 1.			
FT	NON_TER			
FT	1			
SEQUENCE	593 AA: 9BA127E19B18C0D8 CRC64;			

	Query Match	42.0%	Score 143.5	DB 10	Length 593
	Best Local Similarity	38.8%	Pred. No. 2.2e-08		
	Matches 31	Conservative 10	Mismatches 16	Indels 23	Gaps
QY	2	QROVQCCGRCQEEQOQGGREQOQCCORKCMDEYKEQ---	ERS-----	EH 41	
	::            ::	:    ::			
Db	120	QOQTHRCORRCQIQEQSPERQROCCQRCERYKEQGGREKPEASPRRESRGREEDQRH		1/9	
QY	42	ENYHNHK---KNRSEEEQO 58			
	::	::	::		
Db	180	NPYTFHSQSIKSRHSESEGE 199			

RESULT	3
Q9SP14	
ID	Q9SP14 PRELIMINARY; PRT; 666 AA.
AC	Q9SP14;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE	VICILIN PRECURSOR.
CN	AMP2.
OS	Macadamia integrifolia (Macadamia nut).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX	NCBI_TaxID=60698;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=NOT KNOWN;
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT	"A family of antimicrobial peptides is produced by processing of a 7S
RL	globulin protein in Macadamia integrifolia.";
PL	Plant J. 0:0-0(1999) .
DR	EMBL; AF161884; AAD54245.1; -;
DR	HSSP; P02853; PPHL.
DR	INTERPRO: IPR001113; -;
DR	Pfam: PF00546; Seedstore_7s; 1.
DQ	SEQUENCE 666 AA; 78243 MW; OECA22F8710F8A7B CRC64;

Query Match	38.2%;	Score 130.5;	DB 10;	Length 666;
Best Local Similarity	34.3%;	Pred. No. 6.6e-07;		

Matches	24: Conservative	12: Mismatches	23: Indels	11: Gaps	1
QY	2	0K0Y00C0G0R0E00G0E00C0RKC-----	W0QY0F0E0R0E0N0Y0N0N0K0N	50	
		:	:		
Db	123	000Y00C0G0R0E0N0T0C00R0C0R0Y0E0K0K0Q0K0Y0E0D0R0D0E0K0Y0E0R0K0E		182	
QY	51	RSEEEG00R	60		
		: :			
Db	183	EDNKKRDP00R	192		

RESULT	4			
OSPL3				
ID	09SPL3	PRELIMINARY;	PRT;	625 AA.
AC	09SPL3			
DT	01-MAY-2000 (TREMBLE)	13, Created		
DT	01-MAY-2000 (TREMBLE)	13, Last sequence update		
DT	01-OCT-2000 (TREMBLE)	15, Last annotation update		
DE	VICILIN PRECURSOR (FRAGMENT).			
GN	AMP2.			
OS	Macadamia integrifolia (Macadamia nut).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Eudicotyledons; Proteaceae; Macadamia.			
OX	NCBI_TaxID=60698;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NUIT KERNEL			
RA	Marcus J.P., Coulter K.C., Green J.L., Manners J.M.;			
RT	"A family of antimicrobial peptides is produced by processing of a 7S			
RL	globulin protein in Macadamia integrifolia."			
DR	Plant J. 0:0-0(1999).			
DR	EMBL; AF161885; AAD54246.1;			
DR	HSSP; P02853; 2PHL.			
DR	INTERPRO; IPR001113;			
DR	PFAM; PF00546; Seedstore_7s; 1.			
FT	NON_TER	1		
SO	SEQUENCE	625 AA;	73586 MW;	415808A89D370296 CRC64;

Query Match Similarity	37.3%	Score 127.5	DB 10	Length 622
Best Local Similarity	34.3%	Pred. No. 1.3e-06		
Matches	24	Conservative	13	Mismatches 22
				Indels 11
				Gaps 1
OY	2	OROVOCGCGDEQDQOQCGDEQOQCRK-----	WQYQKDEKREKENEYNNKKKN	50
DB	82	OOOIOOCCKCRKRETPRNMDICQCRKRYEKERKQCKQCKQCKRDEKDEKRYERKKE		141
OY	51	RSEEEGQQR	60	
DB	142	GDNRKDPQR	151	

RESULT	5			
09SPL5				
ID	09SPL5	PRELIMINARY;	PRT;	666 AA.
AC	09SPL5			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VICILIN PRECURSOR.			
GN	AMP2.			
OS	Macadamia integrifolia (Macadamia nult).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Manoilophyta; eudicotyledons; Proteaceae; Macadamia.			
OX	NCBI_TaxID=60698;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NUIT KERNEL			
RA	Marcus J.P., Goulet K.C., Green J.L., Manners J.M.;			
RT	"A family of antimicrobial peptides is produced by processing of a 7S			
RT	globulin protein in Macadamia integrifolia kernels.";			
RL	Plant J. 0:0-0(1999).			
EMBL	AF161883; AAD54244.1; -.			



01-JAN-1998 (TREMBlurel\_05, Last sequence update)  
 DT 01-NOV-1998 (TREMBlurel\_08, Last annotation update)  
 DE CACH44 (FRAGMENT).  
 GN CACH44.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=BRAIN;  
 RC MEDLINE=97369492; PubMed=9225980;  
 RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.H., Kidwai A.S.,  
 RA Breschel T.S., Scline O.C., Gallahan C., McInnis M.G., Ross C.A.;  
 RT "CNAs with Long CG Trinucleotide repeats from human brain."  
 RL Hum. Genet. 100:114-122(1997).  
 RN [2]  
 RP SEQUENCE OF 1-86 FROM N.A.  
 RA Mixx, P., Hinds, K., Suterer, C., Becker, M., Ozerzky P.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-86 FROM N.A.  
 RA Waterston R.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U80741; AAB91439.1; -  
 DR EMBL; AC003992; AAB96326.1; -  
 FT NON TER 304 304  
 SE SEQUENCE 304 AA; 34400 MW; DFB9D9E6BB877A1E CRC64;

```

Query Match 29.7%; Score 101.5; DB 4; Length 304;
Best Local Similarity 35.0%; Pred No. 0.00054;
Matches 21; Conservative 20; Mismatches 16; Indels 3; Gaps 1

QY 1 LQRTQCCGRCQEDGGGREGDCCCRKCMEDYKEDGEGEHENYHNKKNRSEEEGQOR 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 151 LQQGQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 207
    ---00000000000000000000000000000000000000000000000000000

```

RESULT	ID	PRELIMINARY;	PRT;	1556 AA.
0906A1	0906A1			
AC	0906A1			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)			
DE	HOMEODOMAIN TRANSCRIPTION FACTOR PROSPERO.			
GN	PROS.			
OS	Drosophila virilis (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.			
OC	Ephyraoidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7244;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Xu C., Kaufmann R.C., Zhang J., Kladny S., Carthew R.W.;			
RT	*Overlapping activators and repressors delimit transcriptional			
RT	response to receptor tyrosine kinase signals in the Drosophila eye".			
RL	Submitted (Sep-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF190405; AF06660.1; --			
DR	Homeobox: DNA-binding; Nuclear protein.			
XQ	SEQUENCE 1556 AA; 171029 MW; 66F6ACEA2D73E644 CRC64;			

	Query Match	29.4%	Score 100.5	DB 5	Length 1566;
	Best Local Similarity	35.5%	Pred. No.	0.0029;	
	Matches	22:	Conservative	19;	Mismatches 18; Indels 3; Gaps 1
QY	2 QROYCCGCGRCGEEQQGGGRQRQQQCARKCEKYEKKEDR---GEHEHTHHNKKNRSEEBEG	58			
Dd	773 0000000000000000000000E00RRTEDEOEORRRMEBEOOOLHROQHQLDQLQOOO	832			
	59 QR 60				

Db 833 1: 834

RESULT	11	
Q9U1K4		
ID	Q9U1K4	PRELIMINARY;
		PRT;
		523 AA

DT 01-MAY-2000 (TREMBL:rel\_13, Created)  
 DT 01-MAY-2000 (TREMBL:rel\_13, Last sequence update)  
 DT 01-OCT-2000 (TREMBL:rel\_15, Last annotation update)  
 DE C2H2 ZINC FINGER TRANSCRIPTION FACTOR.  
 GN D-Sp1.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON S.  
 RA SCHOCK F., PARNELL B.A., WIMMER E.A., JACKLE H.;  
 RT "Common and diverged functions of the Drosophila gene pair D-Sp1 and  
 buttonhead";  
 RL Mech. Dev. 89:125-135(1999).  
 DR EMBL: AJ131022; CAB55429.1; -.  
 DR HSSP: P08047.1SP2.  
 DR INTERPRO: IPR000822; -.  
 DR PFAM: PF00006; zf-C2H2.3  
 DR PROSITE, PS00028; ZINC\_FINGER\_C2H2.3.  
 SO SEQUENCE 523 AA; 54577 MW; 64FB8392A72A7AB8 CRC64;

```

Query Match      29.2%; Score 100; DB 5; Length 523;
Best Local Similarity 33.9%; Pred. No. 0.0013;
Matches 20; Conservative 18; Mismatches 19; Indels 2; Gaps 1;

OY 2 OROYOCCGRCCEODGCGREDDGCCRCMEQCKEEDRGCHENYNHNHKKNRSEEGSCOR 60
bb ::::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
429 -OOOHOOOOOOOOOOOOHHOOOOOOOH--OOOOOOHHOOOOOOHHOOOOOOHHOOOOOOO 485
```

RESULT	12	
09W318		
ID	09W318	PRELIMINARY;
AC	09W318;	PRT; 666 AA.
DT	01-MAY-2000 (TREMBlrel. 13, Created)	
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)	
DT	01-OCT-2000 (TREMBlrel. 15, Last annotation update)	
DE	SP1 PROTEIN.	
GN	Sp1.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_Taxid=7227;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BERKELEY;	
RX	MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Amaratilakis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA	Georgakilas S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,	
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Arrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beeson K.Y., Beus P.V., Berman B.P., Bhandari D., Bolshakov S.,	
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,	
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	

RA Dodson K., Doup L.E., Downes M., Dugan-Recha S., Dunkov B.C., Dunn P.  
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
CA Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekhlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paole J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson R., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Teector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh K.-F., Zaveri J.S., Zhai M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000)  
DR EMBL: AE003448; AAF46519.1; -;  
DR HSP: P08047; ISP2.  
DR PLAYBASE: FBgn0020378; Spl.  
DR INTERPRO: IPR000832;  
DR PRAM; PF000936; ZF-C2H2; 3.  
DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 3.  
QO SEQUENCE 666 AA; 67658 MW; 9DAFAFB2156D1148 CRC64;

	Query Match	29.2%	Score 100	DB 5	Length 666
	Best Local Similarity	31.7%	Pred No. 0	0016	
	Matches 19;	Conservative 19;	Mismatches 22;	Indels 0;	Gaps 0.
Oy	1 LORQYOCGRCEODGCGREDDGCCKKCMEDYKDEDERGENHNNHKKNRSEEGDQR	60			
	: :	:	:	:	:
Db	LHQOQQQQQQHHQQQQQQQQQQQQHHQQQQQQQQQQQQHHQQQQQQQQHHQQQQQQQQ	628			
	: :	:	:	:	:

RESULT	13	
061118		
ID	061118	PRELIMINARY; PRF: 171 AA.
AC	061118;	
DT	01-NOV-1996 (TREMBLrel. 01. Created)	
DT	01-NOV-1996 (TREMBLrel. 01. Last sequence update)	
DT	01-NOV-1999 (TREMBLrel. 12. Last annotation update)	
DE	GLUTAMINE REPEAT PROTEIN 1. GLRP1.	
GN		
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C3H/HEJ; TISSUE=BONE MARROW;	
RX	MEDLINE=96411770; Pubmed=8810323;	
RA	Cox G.W., Taylor L.S., Willis J.D., Melillo G., White R.L.,	
RA	Anderson S.K., Lin J.J.;	
RT	"Molecular cloning and characterization of a novel mouse macrophage	
RT	gene that encodes a nuclear protein comprising polyglutamine repeats	
RL	and interspersing histidines.";	
RT	J. Biol. Chem. 271:25515-25522(1996).	
DR	EMBL: U46463; AAC52934.1; -.	
DR	KCD; MG1:108038; G1RP1.	
SQ	SEQUENCE 171 AA; 21572 MW; C6A117E6FE26DE5 CRC64;	

Query Match	28.1%;	Score 96;	DB 11;	Length 171;
Best Local Similarity	30.2%;	Pred. No. 0.0013;		

[illegible]

```

Query Match      27.6% Score 94.5; DB 11; Length 420;
Best Local Similarity 36.6% Pred. No. 0.0042;
Matches 19; Conservative 20; Mismatches 18; Indels 5; Gaps 1;

QY    4 QYQQGCGCQDQGGGGRQRDQC-----QRKCMGYTKEDRGEGNEHYNNHKKNKRSFEESQ 58
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     173 QQQQQQQQQQQQQQQQQQQQQQQQQQQQRFNHNHQKQDFNHNHQQQQQQFNDNNHHNQO 232
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     59 QR 60
      | :
Db     233 QQ 234

```

RESULT	15		
Q9VZU6			
ID	Q9VZU6	PRELIMINARY;	PRT; 828 AA.
AC	Q9VZU6: Q24005:		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)		
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)		
DE	BRVVI1 PROTEIN (BRV-VII PROTEIN DOMAIN).		
GN	BRVVI1.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Phytophidae; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRATIN-BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans R.A., Gocayne J.D.,		
RA	Manantiles P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,		

